Piano Fundamental Frequency Estimation Algorithm Based on Weighted Least Square Method

Shi Lixin, Zhang Junxing, Pang Lingbin
Dept. of Electromechanical and Information Engineering, Dalian Nationalities University, Dalian 116600, China
shilixin@dlnu.edu.cn; zhangjunxing@dlnu.edu.cn, plingb@dlnu.edu.cn

ABSTRACT: Fundamental frequency estimation(FFE) has been studied for many years, and current techniques are still not accurate and robust. Because of the fundamental frequency losing, the harmonic losing and the inharmonicity of piano sound, these FFE methods are invalid generally. The mean square error of the harmonic frequency based on quadratic interpolation of FFT is discussed, and the influence of other harmonic sidelobe on the harmonic measurement with the different length Hanning, Hamming, Chebyshev and Blackmanharris windows is analyzed. A FFE algorithm using weighted least square method is derived. The simulation experiment shows that the method is feasible, and the relative error is about 10^{-5}.

KEYWORDS: fundamental frequency estimation; weighted least square method; harmonic structure; influence of sidelobe

I. INTRODUCTION

Fundamental frequency estimation has been a popular research topic for many years, and many estimation algorithms have been proposed. Kedem[1] detected the periodicity by means of the zero-crossing method. Cheveigné and Kawahara[2] employed a cumulative mean function to improve auto-correlation approach. Engdegard[3] extended the envelope periodicity method by considering both negative and positive amplitudes of the envelope. Piszczalski and Galler[4] predicted musical pitch from component frequency ratios. Dorken and Nawab[5] improved musical pitch tracking using principal decomposition analysis. Moorer[6] applied optimum comb filter on the transcription of musical sound, which is computationally intensive. Raphael[7] proposed a pitch detector based on the Hidden Markov Model. Doval and Rodet[8,9] presented a series of papers on f_0 estimation using maximum likelihood approach. Most of these methods focus on detecting the existence of notes in the samples instead of measuring the frequency, and concern about the note error rate instead of the accuracy. Each algorithm has been designed for a particular research problem and might not be appropriate for other problem. Unfortunately, there is not any algorithm designed for the particularity of piano spectrum.

Piano is one kind of attacking string instruments. The sound generated by strings seems to be composed of a fundamental frequency which is with the lowest frequency and the biggest amplitude, and many harmonics which are whole-number times as high as the fundamental frequency. However, because of the complex structure of piano and many other factors, the sound does not strictly follow the simple rules as we described above. There are some particularities. (i) The fundamental frequency losing: The amplitude of fundamental frequency is much smaller than harmonic's in many lower notes, and is even zero in several lowest notes. (ii) The harmonic losing: There is the only fundamental frequency in higher notes. The amplitude of harmonic is very small and even nonexistent. (iii) The inharmonicity: The frequency of harmonic is not whole-number times as high as fundamental frequency. The interval between two harmonics becomes bigger gradually with the increasing of harmonic order in spectrum. The equation of the harmonic[10] is:

$$f_k = kf_0\sqrt{1+(k^2-1)B} \tag{1}$$

where f_0 is the fundamental frequency, k is the harmonic order and B is the inharmonicity coefficient of the note, which is relevant to the instrument and note. Equation (1) can be written in logarithmic form as:

$$\ln f_k = \ln k + \ln f_0 + \frac{1}{2}\ln[1+(k^2-1)B] \tag{2}$$

The particularities of harmonic structure affect the feasibility and accuracy of common FFE methods. Piano waveforms with rich harmonic may cross zero many times in one period. It is difficult to determine which peak is the fundamental frequency and which peak represents harmonic based on auto-correlation techniques. The phenomenon of the fundamental frequency losing makes harmonic peak approach fail because a higher harmonic is wrongly regarded as the fundamental frequency. The phenomenon of the inharmonicity makes optimum comb filter algorithm produce a bigger error. In this paper, we study the particularities of piano spectrum fully, estimate harmonics using quadratic interpolation of FFT, and calculate the fundamental frequency based on weighted least square(WLS) algorithm.

II. FFE ALGORITHM USING WLS METHOD

A signal of piano note can be expressed as: $s(t) = \sum_i a_i \sin(2\pi f_i t + \phi_i) + n(t)$, where a_i, f_i, ϕ_i are the amplitude, the frequency and the phase respectively, and $n(t)$ is noise. The number of harmonics is variable with the piano and note, so as the parameters of harmonics.

Rife and Vincent[11] enhanced the accuracy of the frequency and the amplitude of the sinusoidal signal using two spectral lines in DFT mainlobe. Jain et al.[12] proposed a sinusoidal signal parameter estimation algorithm based on quadratic interpolation FFT. Abe and Smith[13] investigated the bias caused by quadratic interpolation, and proposed the simple bias correction functions. These approaches which

978-0-7695-3600-2/09 $25.00 © 2009 IEEE
DOI 10.1109/IFITA.2009.402

have enhanced the estimation accuracy as well as reserved the high speed of calculation advantage of FFT are applied widely. Therefore, we can estimate the frequency and the amplitude of each harmonic by quadratic interpolation FFT algorithm.

A. Harmonic estimation error

QI and JIA[14] theoretically analyzed the estimation accuracy of quadratic interpolation FFT algorithm, and simulated it using computer. In the noisy environment, the MSE of δ , which is the MSE of frequency, can be expressed as:

$$Var(\delta) = mse(\delta_\alpha) + mse(\delta_e) \quad (3)$$

where $mse(\delta_\alpha)$ is the MSE of frequency produced by the error of two spectral lines in DFT mainlobe, and $mse(\delta_e)$ is the MSE of frequency produced by the wrong second maximum spectral line in DFT mainlobe. In general, $mse(\delta_e)$ is much smaller than $mse(\delta_\alpha)$ .When any window is applied, $mse(\delta_e)$ can be neglected, and $Var(\delta)$ can be decided by $mse(\delta_\alpha)$ only. $mse(\delta_\alpha)$ is achieved according to the following rule:

$$Var(\delta_\alpha) = \frac{f(\delta, N)}{SNR} \quad (4)$$

Equation (4) indicates that $Var(\delta_\alpha)$ is the reciprocal ratio of *SNR*. $Var(\delta_\alpha)$ becomes smaller as *SNR* gets bigger. $Var(\delta_\alpha)$ is generally relevant to δ . When $\delta < \delta_{th}$, $f(\delta, N)$ changes remarkably without a window, and $f(\delta, N)$ is relatively smooth and nearly reaches a constant with a window.

B. Influence of sidelobes

The accuracy of estimating poly-harmonic signal using quadratic interpolation FFT algorithm is not analyzed in [14], but only the sinusoidal signal. Therefore, we should consider the influence produced by other harmonics.

Fig. 1 and Fig. 2 show the influence of the sidelobes of A_2 note, where longitudinal coordinate is the sum of leakages of all harmonics. The influence of sidelobes is relevant to the window function according to the experimental results. (i) If the sidelobes decay slowly, the influence is proportionate. Because all the sidelobes of Chebyshev window have the same height, the influence of the sidelobes of Chebyshev window is the most proportionate. (ii) If the sidelobes decay rapidly, the influence becomes weaker when the window gets longer. In case of Hanning window, the influence of sidelobes is -87dB when the length of window is 4096, and the influence is -141dB when the length is 32768. Thus, the influence of sidelobes can be treated as noise when a proportionate window is selected, and can be neglected when the sum of leakages of all harmonics is much smaller than noise.

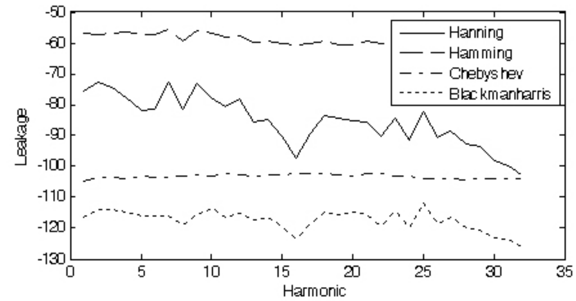

Figure 1. Influence of sidelobes(*N*=4096)

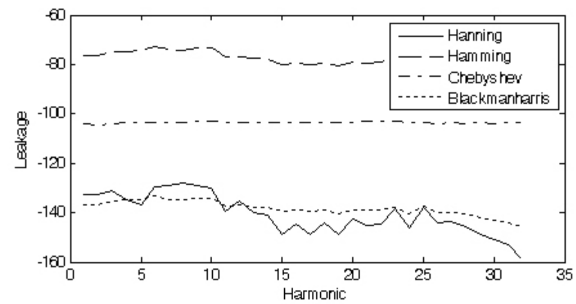

Figure 2. Influence of sidelobes(*N*=32768)

C. Weighted coefficient

In case of Gaussian white noise, the MSE is σ_n^2 , and the signal noise ratio is $SNR = a^2/(2\sigma_n^2)$. When a proportionate window is selected, the equation can be re-written as:

$$\mathrm{var}(\tilde{f}) \approx \frac{C}{SNR} = \frac{2\sigma_n^2 C}{a^2} \quad (5)$$

where *C* is constant. Then, we have:

$$\mathrm{var}(\ln \tilde{f}) = \frac{\mathrm{var}(\tilde{f})}{\tilde{f}^2} = \frac{2\sigma_n^2 C}{\tilde{f}^2 a^2} \quad (6)$$

Because the frequency and the amplitude of every harmonic are different, the MSE of every harmonic is different, too. Thus, the weighted coefficient is defined as :

$$\beta_i = f_i^2 a_i^2 \quad (7)$$

Equation (7) indicates that the harmonic with bigger amplitude makes important contribution to FFE, so does the higher order harmonic.

D. WLS algorithm

We assume that the estimation value of the fundamental frequency is $\tilde{f}_0$, and the estimation value of the inharmonicity coefficient is $\tilde{B}$. According to (2), the estimation value of the harmonic is

$\ln \widetilde{f}_k = \ln k + \ln \widetilde{f}_0 + \frac{1}{2}\ln[1+(k^2-1)\widetilde{B}]$. The square error of the harmonic is given by $\varepsilon_k = (\ln \widetilde{f}_k - \ln f_k)^2$.

In order to calculate easily, we define $\widetilde{F} = \ln \widetilde{f}_0$. Meanwhile, we deal with $\ln[1+(k^2-1)\widetilde{B}]$ using Taylor's expansion. As neglecting the higher order, we can have $\ln[1+(k^2-1)\widetilde{B}] \approx (k^2-1)\widetilde{B}$.

The square error can be calculated as: $\varepsilon = \sum_k \beta_k (\ln \widetilde{f}_k - \ln f_k)^2$. When $\frac{\partial \varepsilon}{\partial \widetilde{F}} = 0$ and $\frac{\partial \varepsilon}{\partial \widetilde{B}} = 0$, the errors of $\widetilde{F}$ and $\widetilde{B}$ are the minimums respectively.

When $\frac{\partial \varepsilon}{\partial \widetilde{F}} = 0$, we have:

$$\sum_k \widetilde{f}_k^2 a_k^2 \widetilde{F} + \sum_k \frac{1}{2} \widetilde{f}_k^2 a_k^2 (k^2-1)\widetilde{B} = \sum_k \widetilde{f}_k^2 a_k^2 \ln f_k - \sum_k \widetilde{f}_k^2 a_k^2 \ln k \tag{8}$$

When $\frac{\partial \varepsilon}{\partial \widetilde{B}} = 0$, we have:

$$\sum_k \widetilde{f}_k^2 a_k^2 (k^2-1)\widetilde{F} + \sum_k \frac{1}{2} \widetilde{f}_k^2 a_k^2 (k^2-1)^2 \widetilde{B} = \sum_k \widetilde{f}_k^2 a_k^2 (k^2-1)\ln f_k - \sum_k \widetilde{f}_k^2 a_k^2 (k^2 1)\ln k \tag{9}$$

Equation (7) and (8) can be expressed in array:

$$A \cdot X = C \tag{10}$$

where
$$A = \begin{pmatrix} \sum_k \widetilde{f}_k^2 a_k^2 & \sum_k \frac{1}{2} \widetilde{f}_k^2 a_k^2 (k^2-1) \\ \sum_k \widetilde{f}_k^2 a_k^2 (k^2-1) & \sum_k \frac{1}{2} \widetilde{f}_k^2 a_k^2 (k^2-1)^2 \end{pmatrix}$$

$$X = \begin{pmatrix} \widetilde{F} \\ \widetilde{B} \end{pmatrix}$$

$$C = \begin{pmatrix} \sum_k \widetilde{f}_k^2 a_k^2 \ln f_k - \sum_k \widetilde{f}_k^2 a_k^2 \ln k \\ \sum_k \widetilde{f}_k^2 a_k^2 (k^2-1)\ln f_k - \sum_k \widetilde{f}_k^2 a_k^2 (k^2-1)\ln k \end{pmatrix}$$

The solution to (10) is given by $X = A^{-1} \cdot C$. Then, we can have $\widetilde{F}$ and $\widetilde{B}$, as well as $\widetilde{f}_0$.

III. SIMULATION EXPERIMENTS

In order to analyze the accuracy conveniently, we extract the harmonic structures of each note of a piano, and calculate the inharmonicity coefficient of each note using the harmonic whose amplitude is bigger than ten percent of the maximum harmonic. Then, the testing signals can be constructed by the harmonic structure, the inharmonicity coefficient, as well as the standard fundamental frequency.

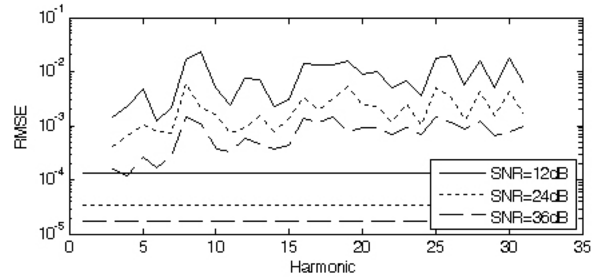

Figure 3. Comparison of root mean square error of C_1 note

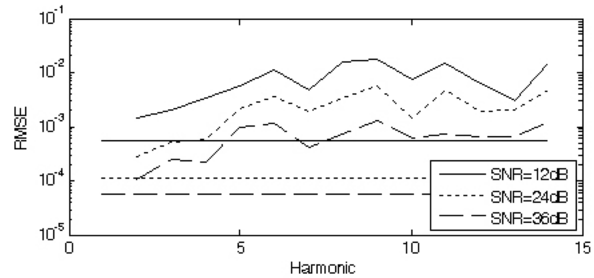

Figure 4. Comparison of root mean square error of C note

A set of simulation experiments has been performed to study the performance of our FFE algorithm using WLS method by Monte-Carlo approach(200 times). Fig. 3 and Fig. 4 show the comparison of the root mean square error, where the curves present the root mean square error (σ_i) of every harmonic of C_1 note and C note, and the straight lines present the root mean square error(σ_0) which is calculated with our FFE algorithm. The simulation experiments show that σ_0 is smaller than σ_i of every harmonic, σ_0 decreases as *SNR* increases, and σ_0 becomes smaller when there are more harmonics.

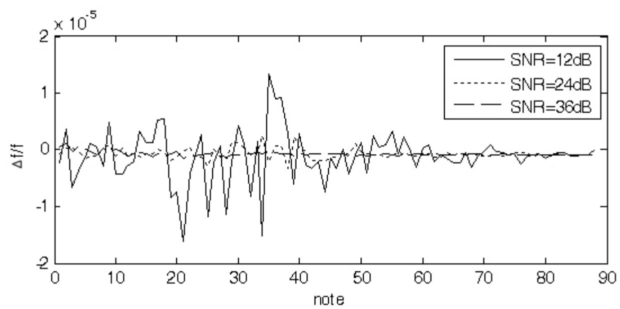

Figure 5. Relative error of fundamental frequency

Fig. 5 shows the relative error of the fundamental frequency of all 88 notes of a piano. The inputs are the signals with Hanning window(N=32768) which are constructed above. The simulation experiments show that the relative error decreases as *SNR* increases, the relative error increases as the number of harmonic decreases, and the relative error is about 10^{-5}.

IV. CONCLUSIONS

The accuracy of traditional FFE methods is very low because these methods neglect the piano sound particularities of the fundamental frequency losing, the harmonic losing and the inharmonicity. In this paper we propose a FFE algorithm using WLS method. Based on the error of the harmonic estimation using quadratic interpolation FFT algorithm and the influence of other harmonic sidelobes, the weighted coefficient equation is derived. The simulation experiments show that the relative error is about 10^{-5} which is much less than a cent. The algorithm meets the demands of adjusting piano intonation completely. As the accuracy of FFE is decided by the accuracy of each harmonic, the algorithm would fail if the harmonic can not be distinguished.

REFERENCES

[1] B.Kedem, "Spectral analysis and discrimination by zero-crossings", Proceedings of the IEEE, vol. 74, Nov. 1986, pp. 1477-1493.

[2] A.Cheveigné, and H.Kawahara, "Yin: A fundamental frequency estimator for speech and music", Journal of the Acoustical Society of America, vol. 111, Apr. 2002, pp.1917-1930.

[3] J.Engdegard, Signal processing for a real time phonetograph, Stockholm:Tal, musik och horsel KTH, 2000.

[4] M.Piszczalski, and B.A.Galler, "Predicting musical pitch from component frequency ratios", Journal of the Acoustical Society of America, vol. 66, Sep. 1979, pp. 710-720.

[5] E.Dorken, and S.H.Nawab, "Improved musical pitch tracking using principal decomposition analysis", In International Conference on Acoustics, Speech and Signal Processing, 1994, pp. 217-220.

[6] J.A.Moorer, "On the transcription of musical sound by computer", Computer Music Journal, 1977, pp.32-38.

[7] C.Raphael, "Automatic transcription of piano music", In International Symposium on Music Information Retrieval (ISMIR), Paris, 2002.

[8] B.Doval, and X.Rodet, "Estimation of fundamental frequency of musical sound signals", In International Conference on Acoustics, Speech and Signal Processing, 1991, pp.3657-3660.

[9] B.Doval, and X.Rodet, "Fundamental frequency estimation and tracking using maximum likelihood harmonic matching and HMMs", In International Conference on Acoustics, Speech and Signal Processing, 1993, pp. 221- 224.

[10] N.H.Fletcher, and T.D.Rossing, The Physics of Musical Instruments (2nd edition), Berlin: Springer-Verlag, 1998.

[11] D.C.Rife, and G.A.Vincent, "Use of the discrete Fourier transform in the measurement of frequencies and levels of tones", The Bell System Technical Journal, 1970, pp. 197- 228.

[12] V.K.Jane, W.L.Collins, and D.Davis, "High-accuracy analog measurements via interpolated FFT", IEEE Transactions on Instrumentation and Measurement, vol. 28, Jun. 1979, pp.113-122.

[13] M.Abe, and J.O.Smith III, "Design Criteria for Simple Sinusoidal Parameter Estimation based on Quadratic Interpolation of FFT Magnitude Peaks", Institute of Electronics, Information and Communication Engineers, Oct. 2004, pp.7-12.

[14] QI G., and JIA X., "Accuracy Analysis of Frequency Estimation of Sinusoid Based on Interpolated FFT", Acta Electronica Sinica, vol. 32, Apr. 2004, pp.625-629.